

PH NL 010110Wb	MAT. DOSSIER
-------------------	-----------------



(19) BUNDESREPUBLIK

DEUTSCHLAND



DEUTSCHES  
PATENTAMT

(12)

Patentschrift

(10)

DE 43 09 957 C 1

(51)

Int. Cl.<sup>5</sup>:

G 06 K 9/64

G 06 F 15/70

H 04 N 17/00

(21)

Aktenzeichen:

P 43 09 957.2-53

(22)

Anmeldetag:

26. 3. 93

(43)

Offenlegungstag:

—

(45)

Veröffentlichungstag

der Patenterteilung: 14. 7. 94

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

(73)

Patentinhaber:

Media Control-Musik-Medien-Analysen GmbH,  
79761 Waldshut-Tiengen, DE

(74)

Vertreter:

von Bülow, T.,  
Dipl.-Ing.Dipl.-Wirtsch.-Ing.Dr.rer.pol., Pat.-Anw.,  
81545 München

(72)

Erfinder:

Schulze, Klaus, Dr.-Ing., 2412 Nüsse, DE

(56)

Für die Beurteilung der Patentfähigkeit  
in Betracht gezogene Druckschriften:

DE 34 15 685 C2

DE-OS 24 14 809

GB 16 03 940

GB 12 79 451

(54)

Verfahren zum Wiedererkennen von unikaten Bildsignalen und Vorrichtung zur Durchführung des Verfahrens

(57)

Das Verfahren zum Wiedererkennen von unikaten Bildsignalen, insbesondere von unikaten Bildfolgen tastet Luminanzwerte vorbestimmter Pixel und vorzugsweise vorbestimmter Gruppen von Pixeln (Cluster) ab, bildet ein Differenzsignal zwischen den Luminanzwerten der vorbestimmten Cluster und aus diesen Differenzwerten eine Signumfunktion. Pro Vollbild entsteht damit ein binärer Merkmalsvektor, dessen Länge der Anzahl der pro Vollbild abgetasteten Cluster entspricht. Für eine vorbestimmte Anzahl aufeinanderfolgender Vollbilder (z. B. 51) wird dann ein Bildsequenz-Merkmalvektor gebildet, der nach jedem Vollbild aktualisiert wird. Nach jedem Vollbild wird der aktualisierte Merkmalsvektor mit einer Vielzahl von in gleicher Weise ermittelten Merkmalsvektoren, die abgespeicherte, wiederzuerkennende Bildfolgen identifizieren, korreliert. Bei einem vorbestimmten Korrelationswert, wird ein Signal "Bildfolge wiedererkannt" erzeugt.

DE 43 09 957 C 1

DE 43 09 957 C 1

Die Erfindung bezieht sich auf ein Verfahren zum Wiedererkennen von unikaten Bildsignalen, insbesondere von Fernsehbildern und eine Vorrichtung zur Durchführung des Verfahrens. Für verschiedene Zwecke wie z. B. statistische Auswertungen, Abrechnungen, etc. ist es wünschenswert, festzustellen, ob ein bestimmtes Bild oder eine bestimmte Bildfolge in einer Sendung vorhanden ist. Beispielsweise möchten Auftraggeber von Werbespots im Fernsehen überprüfen, ob und wann der Werbespot von einem Fernsehsender abgestrahlt wurde und auch, ob er in voller Länge abgestrahlt wurde oder gekürzt wurde. Ähnliches gilt für Spielfilme und sonstige Fernsehsendungen. Bisher wurde dies entweder von Personen überprüft, die das vollständige Programm überwachten und entsprechende Aufzeichnungen machten oder dadurch, daß das vollständige Programm auf Videorecordern aufgezeichnet wurde und anschließend die Abstrahlung bestimmter Sendungen darauf überprüft wurde. Diese Methoden sind sehr personalaufwendig und aufgrund menschlichen Versagens, Unaufmerksamkeit, etc. nicht zuverlässig genug.

Eine ähnliche Problematik stellt sich auch bei der Überwachung von Sprach- und Musiksendungen, wo dieses Problem in der EP 0 296 588 A2 dadurch gelöst wird, daß von den zu untersuchenden Tonsignalen jeweils Hüllkurven gebildet und diese Hüllkurven auf einen engen, niedrigen Frequenzbereich bandbegrenzt werden. Zeitliche Ausschnitte werden nach Digitalisierung in einem Speicher als "Masken" gespeichert. Bei der Überwachung von Sendungen werden dann die bandbegrenzten Hüllkurvensignale der zu untersuchenden Tonsignale in gleichlangen Abschnitten in einem zweiten Speicher zwischengespeichert. Innerhalb des Zeitraumes zwischen dem Eintreffen zweier aufeinanderfolgender zu untersuchender Signale wird dann der Inhalt des Zwischenspeichers mit allen Masken des ersten Speichers verglichen, wobei bei einem vorgegebenen Grad der Übereinstimmung ein Erkennungssignal erzeugt wird.

Eine naheliegende Lösung für die Überwachung von Fernsehsendungen wäre es, mit diesem aus der EP 0 296 588 bekannten Verfahren, die Tonsignale der Fernsehfilme, Werbespots, etc. auszuwerten. Dies führt aber nicht zu dem gewünschten Erfolg, da beispielsweise manche Werbespots, bei gleicher Bildfolge, regional mit anderem Text oder anderer Musik abgestrahlt werden.

Somit ist es erforderlich, das Bild oder die Bildfolge selbst auszuwerten und zu überprüfen, ob eine abgestrahlte Bildfolge mit einer Referenzbildfolge übereinstimmt.

Die DE-OS 24 14 809 zeigt einen Mustererkenner, bei dem das Bild in Mikrobereiche, d. h. eine Anzahl  $n$  von Pixeln aufgeteilt wird. Zur Komprimierung der Information werden Helligkeitsänderungsvektoren bezüglich der Mikrobereiche bestimmt. Für eine Mustererkennung ist zusätzlich ein Vergleich zum Vergleichen mit mehreren voreingestellten Einheitsvektoren vorgesehen. Ein Vergleichsentscheider vergleicht den erzeugten Eigenschaftscode für jeden Mikrobereich mit einem gespeicherten Eigenschaftscode und erzeugt bei entsprechender Übereinstimmung ein Ausgangssignal.

Verfahren zur Bilderkennung werden in verschiedenen Bereichen der Technik angewandt, beispielsweise bei der industriellen Fertigung, um festzustellen, ob ein bestimmter Gegenstand in einer bestimmten Ausrich-

tung auf einem Förderband transportiert wird, vgl. DE-A1-39 13 620, DE-A1-34 34 624, DE-A1 41 23 713, DE A1 40 34 742, DE-C2-30 01 588 oder DE-OS 24 26 803. All diese bekannten Verfahren untersuchen einzelne Bilder aber keine Bildfolgen. Wollte man mit diesen Verfahren Bildfolgen von Fernsehbildern von einer Wiederholfrequenz von 25 Vollbildern pro Sekunde anwenden, so würde der Aufwand unverträglich ansteigen.

Aus der DE-C2-34 15 685 ist es bekannt, einen Szenenwechsel bei der Film- und Videoproduktion dadurch automatisch zu erkennen, daß eine Szene daraufhin überwacht wird, ob Änderungen der Bildsignale auftreten, die auf ein sich innerhalb der Szene bewegendes Objekt zurückzuführen sind. Das Problem wird dort so gelöst, daß mittlere Helligkeitswerte eines Bildes gebildet und mit den mittleren Helligkeitswerten eines unmittelbar darauf folgenden Bildes laufend verglichen werden, wobei durch Bewegungen verursachte Helligkeitsänderungen durch nachgeschaltete Filter eliminiert werden, derart, daß  $n$  diskrete Differenzwerte auf einen einzelnen Szenenwechselimpuls untersucht werden. Die aus dem Vergleich gewonnenen Ergebnisse werden dann automatisch bewertet und ausgegeben. Für die Wiedererkennung von als Unikate vorliegenden Referenz-Bildfolgen ist dieses Verfahren nicht geeignet, zumal gerade bei Werbesendungen über relativ lange Zeiträume von einigen Sekunden auch unbewegte Standbilder gesendet werden und Szenenwechsel für das Wiedererkennen ganz bestimmter (unikater) Bildfolgen nicht aussagekräftig sind.

Aus der GB-PS 1 603 940 ist ein Kamera-Überwachungssystem bekannt, das automatisch feststellt, ob sich im Kamerabereich befindliche Objekte bewegen. Hierzu werden Helligkeitswerte von mindestens zwei unterschiedlichen Bereichen der Szene abgetastet, in elektrische Signale umgewandelt und mit entsprechenden Werten einer nachfolgenden Abtastung verglichen. Ändern sich die Helligkeitswerte an identischen Bildbereichen innerhalb kurz aufeinanderfolgender Zeitintervalle, so ist dies ein sicheres Zeichen für ein sich bewegendes Objekt im Kamerabereich. Der gleiche Grundgedanke wird ebenfalls bei einem Überwachungssystem gemäß der GB-PS 12 79 451 angewandt. Für die Zwecke der vorliegenden Erfindung ist dieses Verfahren ebenfalls nicht anwendbar, da länger dauernde Standbilder nicht erkannt werden können und da Helligkeitswechsel in aufeinanderfolgenden Bildern nicht genug Informationen enthalten, um als Unikat vorliegende Bildfolgen einwandfrei wiedererkennen zu können.

Aufgabe der Erfindung ist es daher, ein Verfahren und eine Schaltungsanordnung der eingangs genannten Art zu finden, das mit hoher Sicherheit das Wiedererkennen von Bildern oder Bildfolgen gestattet. Ein weiteres Ziel der Erfindung liegt darin, das Verfahren weitestgehend in Echtzeit durchzuführen und auch zu überwachen, ob eine von vielen Bildfolgen aktuell abgestrahlt wird. Dabei soll der technische Aufwand möglichst gering sein, was bedingt, die Datenmenge möglichst zu reduzieren, was wiederum Voraussetzung dafür ist, das Verfahren weitestgehend in Echtzeit ablaufen zu lassen.

Diese Aufgabe wird durch die im Patentanspruch 1 angegebenen Merkmale gelöst. Vorteilhafte Ausgestaltung und Weiterbildungen der Erfindung sind den Unteransprüchen zu entnehmen.

Für die Erkennung eines unikaten Bildes ist es mit Sicherheit ausreichend, alle Bildpunkte eines abgestrahlten Bildes mit allen Bildpunkten eines Referenzbildes zu vergleichen. Dies würde jedoch zu einer zu ver-

gleichenden Datenmenge führen, die in einer Echtzeitverarbeitung nicht mehr zu bewältigen wäre. Ein Vollbild einer Fernsehstrahlung mit beispielsweise 620 Spalten und 512 Zeilen enthält 317.440 Pixel. Würde man hierbei die Luminanzwerte (Helligkeitswerte) auswerten und die Farbinformation unberücksichtigt lassen und mit einer Amplitudenauflösung der Luminanzwerte von 8 Bit arbeiten, was 256 Graustufen entspricht, so hätte man pro Vollbild eine Datenmenge von rund 317 Kbyte pro Bild. Bei 25 Vollbildern pro Sekunde entspräche dies einem zu verarbeitenden Datenfluß von rund 8 Mbyte pro Sekunde. Ein praktikabler und weitestgehend in Echtzeit arbeitender Ansatz zur Erkennung unikativer Bildsignale muß also versuchen, mit Teilmengen dieses großen Datenumfanges auszukommen. Theoretisch genügt schon ein einzelnes Bildelement (Pixel) um bei störungsfreier Übertragung das gesuchte Bild wieder zu finden. Dabei wird man allerdings eine große Anzahl anderer Bilder identifizieren, die zufällig in diesem einen Pixel mit dem gesuchten Bild übereinstimmen (Falscherkennung). Ein Problem besteht also darin, eine möglichst geringe aber doch ausreichende Zahl von Bildmerkmalen zu definieren, die das Maß der Falscherkennung auf ein für den jeweiligen Anwendungszweck vertretbares Maß reduzieren.

Bei der Erfindung werden die Luminanzwerte (Helligkeitswerte) verschiedener Pixel erfaßt. Ein einziger Pixel ist wegen des immer vorhandenen Rauschens für einen Vergleich aber nicht geeignet. Wird beispielsweise der Luminanzbereich von 0 V bis 1 V mit 8 Bit digitalisiert codiert, so repräsentiert eine Quantisierungsstufe der Digitalisierung einen Spannungsbereich von ca. 4 mV. Bei einem betrieblichen Störabstand von 40 dB bezogen auf einen 700 mV Weißwert beträgt die effektive Störampplitude 7 mV und ist damit doppelt so hoch wie eine Quantisierungsstufe. Der Luminanzwert eines Pixels schwankt somit zufällig um mehrere Quantisierungsstufen. Beim Vergleich zweier entsprechender Pixel müßte man ein mehrere Stufen umfassendes Toleranzfenster zugrundelegen. Zusätzlich sprechen gegen die Verwendung eines einzelnen Pixels als Vergleichsmerkmal folgende Gründe: Bei der Bildspeicherung, -übertragung und der -erfassung kann es je nach verwendetem Verfahren zu zeitlichen Ungenauigkeiten kommen, die sich in einer zufälligen Positionsänderung des Pixels, bezogen auf die Spalten oder Zeilen des Bildes, bemerkbar machen. Diese zeitliche Ungenauigkeit kann im Bereich von einigen Nanosekunden liegen. Es würden dann Pixel miteinander verglichen, die in Wirklichkeit keinen Bezug zueinander haben.

Um diese Schwierigkeiten zu umgehen, werden bei der Erfindung sogenannte Cluster gebildet, die aus mehreren Pixeln bestehen. Dabei wird der mittlere Luminanzwert der Pixel eines Clusters oder die Summe der Luminanzwerte der Pixel eines Clusters bestimmt. Vorzugsweise sind die Cluster aus einem räumlich zusammenhängenden Bereich des Bildes entnommen und wiederum hier vorzugsweise aus einem quadratischen Bereich.

Für die Auswahl der Größe der Cluster bei der Erfindung waren folgende Überlegungen maßgeblich: Statistische Messungen an tausenden von Bildern haben gezeigt, daß der von der Norm vorgesehene Spektralbereich für Videosignale von 5 Mhz bei weitem nicht ausgenutzt wird. Demnach bestehen Bilder im Mittel überwiegend aus tieffrequenten Informationen, was auch aus dem Bildinhalt verständlich wird, wo typische Objekte mit bestimmter Mindestausdehnung dargestellt werden.

Es zeigte sich allerdings auch, daß es keine bevorzugte Objektgröße gibt, sondern einen gleitenden Übergang von feinsten Struktur zu groben, homogenen Flächen. Meßtechnisch kann man diese Bildeigenschaften durch die Erfassung der zeitlich gemittelten Varianz bzw. der Standardabweichung eines Zeilenclusters beschreiben. Ein völlig homogenes Gebiet konstanter Luminanz würde die Varianz 0 aufweisen; ein für den gesamten Bildinhalt repräsentatives Cluster hätte die größtmögliche Varianz.

Fig. 1 zeigt die Standardabweichung eines Spielfilms (Werbung) in Abhängigkeit von der Clustergröße und zum anderen die gleiche Messung an bandbegrenztem Rauschen (6 Mhz) als Videosignal. Für die Wahl der Clustergröße gibt es nach diesen Messungen keine eindeutigen, scharf abgrenzbaren Kriterien. Sie sollte gerade so groß sein, daß die genannten Zeitprobleme unterdrückt und der Einfluß des Rauschens gemindert wird. Sie sollten aber nicht so groß sein, daß relevante Bilddetails durch die Mittelwertbildung nicht mehr erkannt werden. Als vernünftiger Kompromiß wird bei einem bevorzugtem Ausführungsbeispiel der Erfindung eine Clustergröße von  $5 \times 5$  Pixeln, angeordnet im Quadrat, gewählt. Diese Wahl ist zunächst jedoch nur als erste Näherung an ein relatives Optimum anzusehen und kann in gewissen Größen geändert werden. Bei dieser Clustergröße wird der Störabstand im günstigsten Fall um den Faktor 5 entsprechend 14 dB durch Mittelung verbessert.

Bei der Erfindung werden mehrere Mittelwerte solcher Cluster miteinander "verglichen". Es stellen sich die Fragen, wie viele Cluster verwendet werden, wie diese in Form von "Masken" auf dem Bild verteilt werden und wie sie zu vergleichen sind. Aufgrund der in Echtzeit zu verarbeitenden Datenmenge, der heutigen Arbeitsgeschwindigkeit von Hardware-Bausteinen und den Genauigkeitsanforderungen an die Bilderkennung kommt die Erfindung zu der Erkenntnis, daß es ausreicht, aus jedem Vollbild von 40 ms Dauer ein Merkmal von 10 Bit abzuleiten. Hierzu schlägt die Erfindung in einem Ausführungsbeispiel vor, 10 Cluster pro Vollbild zu verwenden und die mittlere Luminanz von in einer vorgegebenen Reihenfolge "benachbarter" Cluster miteinander zu vergleichen. Der Vergleich besteht hier in einer Differenzbildung der Luminanzwerte. Aus dieser Differenz wird dann das mathematische Vorzeichen (Signumfunktion) digital codiert. Mit anderen Worten wird bei dem Vergleich nur untersucht, welcher von zwei miteinander zu vergleichenden Clustern heller ist, d. h. einen größeren mittleren Luminanzwert hat. Um auch den Fall zu erfassen, daß zwei "benachbarte" Cluster innerhalb der bei der Digitalisierung auftretenden Quantisierung gleiche Helligkeitswerte haben, wird der Vergleich wie folgt durchgeführt:

$$\begin{aligned} \text{Cluster}(n) < \text{Cluster}(n+1) &= 1 \\ \text{Cluster}(n) > &= \text{Cluster}(n+1) = 0. \end{aligned}$$

Zusätzlich kann für den Vergleich noch eine Hysterese vorgesehen sein, dergestalt, daß die erste Bedingung " $\text{Cluster}(n) < \text{Cluster}(n+1)$ " nur dann als erfüllt angesehen wird, wenn die Differenz der Luminanzwerte der beiden zu vergleichenden Cluster größer als ein vorgegebener Schwellwert ist, während für alle anderen Fälle die zweite Bedingung " $\text{Cluster}(n) > = \text{Cluster}(n+1)$ " als erfüllt angesehen wird. Der Schwellwert kann beispielsweise zwei Digitalisierungsstufen betragen.

Bei zehn Clustern je Bild ergeben sich somit zehn

Differenzen, die sich nach der entsprechenden Vorzeichencodierung als 10 Bit darstellen lassen. Bei diesem Vergleich wird der zehnte Cluster eines Vollbildes mit dem ersten Cluster des nächsten Vollbildes verglichen. Es wäre aber auch möglich, den zehnten Cluster eines Vollbildes mit dem ersten Cluster desselben Vollbildes zu vergleichen. Diese Vorzeichencodierung (Signumfunktion) hat nicht nur den Vorteil einer enormen Datenreduktion, sondern auch, daß der Code unempfindlich gegenüber Pegelschwankungen der Luminanzwerte ist.

Die Lage der Cluster und die Reihenfolge ihrer Abfrage ergibt sich indirekt aus den Überlegungen zur Clustergröße. Während bei der Clusterbildung möglichst korrelierte Bildelemente, d. h. Pixel mit innerer Bindung der Bildbereiche zu einem Cluster zusammengefaßt wurden, müssen für die Lage der Cluster und die Reihenfolge ihrer Abfrage die einzelnen Cluster zur Verbesserung des Informationsgehaltes möglichst unabhängige Bildbereiche repräsentieren. Ihre Abfragerihenfolge wird daher so gewählt, daß sich im Mittel eine möglichst große Sprungweite ergibt. Der Abfrageweg wird dann maximal. Die Cluster können statistisch im Bild angeordnet werden und sind vorzugsweise möglichst gleich über den gesamten Bildbereich verteilt.

Da pro Vollbild nur eine sehr kleine Anzahl, bei 10 Clustern pro Vollbild also nur 10 Bit verwendet werden, um ein Bild zu identifizieren, ist die Wahrscheinlichkeit einer Falscherkennung noch recht hoch. Mit 10 Bit ließen sich theoretisch nur  $2^{10} = 1024$  verschiedene Bilder identifizieren. Jedes nur denkbare Bild ließe sich dann eindeutig einer Klasse von 1024 möglichen Bildern zuordnen. Für die Wiedererkennung unikater Bildsignale ist damit aber ein Merkmalsvektor von 10 Bit zu kurz.

Die Erfindung schlägt daher weiter vor, einen längeren Merkmalsvektor zu bilden, durch entsprechendes Abtasten mehrerer aufeinanderfolgender Bilder. In einem Ausführungsbeispiel der Erfindung werden 51 aufeinanderfolgende Bilder ausgewertet, was bei 40 ms pro Vollbild dann einer Bildfolge von 2,04 Sekunden entspricht. Man erhält dann bei 51 Bildern à 10 Bit einen Merkmalsvektor von 510 Bit.

Bei der beschriebenen Form des Datenformates mit 51 Bildern erhält man sehr gute Erkennungsergebnisse mit Übereinstimmungen von über 90%. Gelegentliche zufällige Übereinstimmungen mit anderen Bildfolgen weisen einen Grad von höchstens 75% auf. Bei einer Erkennungsschwelle von ca. 85% betrug die Erkennungsrate in Laborversuchen mit einem Gerät nach der Erfindung 100% bei 0% Falscherkennungen. Dies galt für Bildsequenzen mit starkem szenischen Wechsel.

Unterscheiden sich die Bilder innerhalb einer Sequenz nicht, beispielsweise wenn mehrere oder alle aufeinanderfolgende Bilder gleich sind, was bei Standbildern der Fall ist, ist der Informationsgehalt des Merkmalsvektors von 510 Bit zu gering und es kommt zu Falscherkennungen. Der Merkmalsvektor von 510 Bit ist ja dann streng periodisch aufgebaut mit einer Periodizität von 10 Bit.

Nach einer Weiterbildung der Erfindung läßt sich der Informationsgehalt jedoch deutlich steigern, indem aufeinanderfolgenden Bildern unterschiedliche Merkmale entnommen werden. Nach diesem Ausführungsbeispiel der Erfindung wird also abwechselnd mit zwei "Masken" für die Anordnung der Cluster gearbeitet. Selbstverständlich liegt es auch im Rahmen der Erfindung mit mehr als zwei solcher Masken zu arbeiten, die dann zyklisch angewandt werden. Dadurch wird selbst bei

Standbildern der Informationsgehalt entsprechend der Anzahl der möglichen Masken verdoppelt oder vervielfacht. Die Clusteranordnungen sind dabei möglichst so zu wählen, daß sie untereinander nicht zu Korrelationen führen, die Clusteranordnungen also orthogonal zueinander sind, d. h. die Anordnungen der Cluster der beiden Masken sind um  $90^\circ$  gegeneinander gedreht.

Bei der Erzeugung des Testmusters und beim Vorgang der Bilderkennung wird also beispielsweise mit zwei Masken gearbeitet, beim ersten Bild mit Maske 1, beim zweiten Bild mit Maske 2, usw., d. h. bei ungeradzahligem Bildern mit Maske 1, bei geradzahligem Bildern mit Maske 2. Beim Korrelationsprozeß, d. h. beim Vergleich der gespeicherten Merkmalsvektoren mit den aktuellen Merkmalsvektoren, ergibt sich das Problem, daß es bei dem zu vergleichenden Bild unbekannt ist, mit welcher der beiden Masken das entsprechende Testmuster codiert wurde. Zur Lösung dieses Problems wird daher jedes Testbild mit beiden möglichen Clusteranordnungen (Masken) codiert und es werden zwei unterschiedliche, zueinander orthogonale Testsequenzen erzeugt und gespeichert. Die Abtastung des laufenden Programmes erfolgt aber nur mit einer Sequenz, so daß auch nur ein laufender Merkmalsvektor gebildet wird. Dieser muß dann auch nicht mit beiden gespeicherten, zueinander orthogonalen Test-Merkmalsvektoren verglichen werden. Die Korrelation erfolgt vielmehr nur mit dem gespeicherten Merkmalsvektor, der mit der gleichen Clusteranordnung (Maske) beginnt, wie das laufende Programm. Wenn beispielsweise grundsätzlich bei allen Testmustern das erste, also ungerade Bild mit der Maske 1 codiert wird, dann wird es nur mit dem Muster des laufenden Programmes korreliert, das zur Zeit ebenfalls mit Maske 1 beim ersten Bild beginnt.

Durch diese orthogonalen Anordnungen von zwei Masken steigt der Korrelationsaufwand nicht, es muß nur sichergestellt sein, daß das jeweils "richtige" Testmuster zur Verfügung steht. Da bei der Abtastung des aktuellen Bildes bekannt ist, mit welcher der beiden Masken die Abtastung begonnen wurde, ist eine eindeutige Zuordnung möglich, so daß dann bei der Korrelation auf denjenigen gespeicherten Merkmalsvektor zugegriffen werden kann, der mit derselben Maske beginnt.

Eine Modifikation der Erfindung sieht vor, daß bei gleich langem Merkmalsvektor für eine Bildsequenz das Verhältnis von Cluster pro Bild und Anzahl der aufeinanderfolgenden Bildern verändert werden kann. Beschränkt man beispielsweise den Merkmalsvektor auf 510 Bit, so können in einem Extremfall 510 Bit pro Vollbild verwendet werden, so daß also nur noch ein Vollbild verglichen wird. Im anderen Extremfall können 510 Vollbilder mit je nur einem Cluster verglichen werden. Dazwischen sind eine Vielzahl von anderen Kombinationen möglich. Anders ausgedrückt kann man grundsätzlich von 510 vorzeichencodierten Clusterdifferenzen sprechen, die bedarfsgemäß gleichmäßig über eine entsprechende Anzahl von Bildern verteilt werden können. Man kann auch noch weiter verallgemeinern indem grundsätzlich bei jeder Testmusterergewinnung von jedem Bild eine Vielzahl von Clusterdifferenzen codiert wird und es wird jedem Referenzmuster ein Schlüsselwort mitgegeben, das angibt, mit welcher Teilmenge des Testmusters korreliert werden soll. Auf diese Weise lassen sich bildgenaue z. B. auf ein Senderlogo abgestellte und szenenorientierte Referenzmuster erzeugen.

Im folgenden wird die Erfindung anhand eines Ausführungsbeispiels im Zusammenhang mit der Zeich-

nung ausführlicher erläutert. Es zeigt

**Fig. 1** Ein Diagramm der Standardabweichung eines Spielfilms in Abhängigkeit von der Clustergröße (Anzahl der Pixel);

**Fig. 2** eine schematische Darstellung der Anordnung der Cluster mit zwei zueinander orthogonalen Masken und das sich hieraus ergebende Referenzmuster;

**Fig. 3** eine schematische Darstellung zur Erläuterung des Korrelationsvorganges bei Verwendung von zwei alternierend wirksamen Masken (Anordnung der Cluster).

**Fig. 4** ein Diagramm der Korrelationsergebnisse bei "richtiger" und "falscher" Korrelation des laufenden Programmes mit den unterschiedlichen Testmustern;

**Fig. 5** ein Blockschaltbild eines Gerätes zur Signalvorverarbeitung für die Erkennung unikater Bildsignale nach einer ersten Variante der Erfindung;

**Fig. 6** ein ähnliches Blockschaltbild für eine zweite Variante nach der Erfindung;

**Fig. 7** ein vereinfachtes Blockschaltbild eines Gerätes zur Durchführung der Korrelation; und

**Fig. 8** ein detaillierteres Blockschaltbild eines Gerätes zur Durchführung der Korrelation nach einem Ausführungsbeispiel der Erfindung.

**Fig. 1** macht deutlich, daß die Standardabweichung von Bildfolgen (hier eines Werbefilmes) mit der Clustergröße ansteigt (Clustergröße = Anzahl der Pixel pro Cluster), so daß es keine eindeutigen, scharf abgrenzbaren Kriterien für die Clustergröße gibt, vielmehr steht es dem Fachmann frei, aufgrund anderer, oben genannter Kriterien, die Clustergröße auszuwählen. Ein dabei zu beachtendes Kriterium liegt im Rauschen des Videosignales, das ab einer Clustergröße von etwa 8 Pixel im wesentlichen konstant bleibt.

**Fig. 2a** zeigt schematisch die Anordnung von zehn Clustern 1–10 bzw. deren Verteilung über die Oberfläche eines Bildes. Durch die Pfeile ist die jeweilige "Reihenfolge" gekennzeichnet, in der die Cluster abgefragt werden, also als (zeitlich) "benachbart" angesehen werden. Zwischen jeweils benachbarten Clustern wird das Vorzeichen der Differenz der Luminanzwerte der Cluster gebildet, also beispielsweise zwischen Cluster 1 und Cluster 2, zwischen Cluster 2 und Cluster 3 usw. bis zwischen Cluster 9 und Cluster 10. Der letzte Cluster (hier Cluster 10) eines Vollbildes wird dann mit dem ersten Cluster (Cluster 1) des nächsten Vollbildes verglichen oder bei einer Variante der Erfindung mit dem ersten Cluster desselben Vollbildes. Die in **Fig. 2a** dargestellte Anordnung und Reihenfolge der Cluster bildet eine erste "Maske".

**Fig. 2b** zeigt eine entsprechende zweite Maske, die gegenüber der ersten Maske der **Fig. 2a** um 90 Grad gedreht ist, zu dieser also orthogonal ist.

Bei der Auswahl "benachbarter" Cluster ist darauf zu achten, daß sie räumlich auf dem Bild einen großen Abstand zueinander haben, damit "benachbarte" Cluster möglichst voneinander unabhängige Bildbereiche repräsentieren. Die Reihenfolge ist daher so festzulegen, daß die räumlichen Abstände "benachbarter" Cluster im Mittelwert möglichst groß sind.

**Fig. 2c** zeigt, wie ein Muster (Merkmalsvektor) durch alternierendes Umschalten zwischen den beiden Masken nach **Fig. 2a** und **Fig. 2b** erzeugt wird. Jeweils ein Vollbild wird mit der ersten Maske nach **Fig. 2a** abgetastet, das nächste Vollbild mit der Maske nach **Fig. 2b**, usw. Im gewählten Ausführungsbeispiel mit einer Sequenz von 51 Vollbildern pro Merkmalsvektor ergibt sich dann das schematisch in **Fig. 2c** gezeigte Muster,

wobei jedes Vollbild 10 Bit des Musters einnimmt.

Wie schon oben erwähnt, besteht jeder Cluster aus mehreren Pixeln, die hier beispielsweise in einem Rechteck von 5 × 5 Pixel angeordnet sind. Die Pixel jedes Clusters sollen möglichst nahe beieinander liegen, um so vom Bildinhalt her möglichst miteinander zu korrelieren, d. h. hinsichtlich ihres Luminanzwertes eine innere Bindung zueinander haben.

**Fig. 3** zeigt schematisch den Vorgang der Korrelation eines Referenzmusters aus einem laufenden Programm mit gespeicherten Testmustern. Zur Erzeugung des Testmusters wird jede Bildfolge (hier 51 Bilder) mit beiden Masken nach **Fig. 2a** und **Fig. 2b** abgetastet. Die mit Maske 1 beginnende Abtastung bildet einen ersten Merkmalsvektor, der in einem Testspeicher 1 abgespeichert ist. Die mit Maske 2 beginnende Abtastung bildet einen zweiten Merkmalsvektor, der in einem Testspeicher 2 abgespeichert wird. Die beiden Merkmalsvektoren sind also zueinander orthogonal. Das laufende Programm (Referenzmuster) wird in einer vorgegebenen und damit bekannten Maskenfolge abgetastet, d. h. entweder beginnend mit Maske 1 oder Maske 2. Wird das Referenzmuster entsprechend **Fig. 2c** beginnend mit Maske 1 erzeugt, so muß der Vergleich mit dem im Testspeicher 1 abgelegten ersten Merkmalsvektor erfolgen. Beginnt die Abtastung dagegen mit der Maske 2 gemäß **Fig. 2b** so erfolgt der Vergleich mit dem im Testspeicher 2 abgelegten Merkmalsvektor.

Da jeweils nach jedem Vollbild die Maske gewechselt wird und nach jedem Vollbild ein Vergleich zwischen dem Testmuster und dem Referenzmuster stattfindet, wird der Vergleich immer abwechselnd mit dem im Testspeicher 1 bzw. dem im Testspeicher 2 abgelegten Merkmalsvektor erfolgen. Pro Referenzmuster muß aber trotzdem jeweils nur ein Vergleich durchgeführt werden.

**Fig. 4** zeigt den Unterschied zwischen einem Vergleich mit dem "richtigen" und einen mit dem "falschen" Testmuster. Bei einem falschen Testmuster würde also der Vergleich im Ergebnis um ein Bild versetzt erfolgen, was wegen der Orthogonalität beider Anordnungen keine Übereinstimmung ergäbe. Die obere Kurve zeigt das für Übereinstimmung typische Maximum. Die Korrelation deutet sich also schon einige Bilder vorher an, erreicht das Maximum und fällt dann in ähnlicher Weise wieder ab. Durch die orthogonalen Anordnungen der beiden Masken steigt der Korrelationsaufwand nicht.

**Fig. 5** zeigt ein Blockschaltbild eines Gerätes für die Signalvorverarbeitung. Das übliche Videosignal eines Fernsehers wird einem Videoeingang 11 zugeführt. Dieses Signal entspricht der CCIR-Norm und ist vom Typ BAS mit einer exakt auf 0 Volt geklemmten Schwarzschar. Die Bildamplitude beträgt 0,7 V. Die Sync-Amplitude –0,3 V. Somit ist dem Videoeingang 11 ein handelsübliches Gerät vorgeschaltet, das den Farbräger beseitigt und die Schwarzschar auf 0 Volt festklemmt. Aus diesem Videosignal wird in einem ersten Baustein 12 eine Taktaufbereitung vorgenommen. Derartige Bausteine sind im Handel erhältlich und extrahieren aus dem Videosignal folgende Synchronsignale: Austastimpuls BL, Zeilensynchronimpuls H, Halbbildsynchronimpuls V und einen Vollbildimpuls FR.

Mit Hilfe dieser Signale werden in einem als Steuerzentrale bezeichneten Block 13 ein Spalten- und ein Zeilenzähler gespeist, mit Hilfe derer Ausgangssignale jedes Pixel eines Fernsehbildes nach Spalte und Zeile eindeutig identifiziert werden kann. Mittels eines als Codierfeld bezeichneten Blockes 14 werden die einzelnen

Cluster, d. h. die zu den einzelnen Clustern zu rechnenden Pixel definiert, so daß dann mit Hilfe der in der Steuerzentrale 13 vorhandenen Zeilen- und Spaltenzähler die jeweiligen "Adressen" der auszuwählenden Pixel bestimmt sind. Die Steuerzentrale 13 steuert Halbleiterschalter  $15_1, 15_2 \dots 15_{n+2}$  an, wobei  $n$  die Anzahl der abzutastenden Cluster pro Vollbild ist. Es sind also zwei Halbleiterschalter mehr vorgesehen als Cluster pro Vollbild abgetastet werden. Jeder Halbleiterschalter ist mit einem Integrator  $16_1, 16_2 \dots 16_{n+2}$  verbunden, derart, daß bei geschlossenem Halbleiterschalter 15 der Videoeingang auf den Eingang des zugeordneten Integrators 16 durchgeschaltet wird. Wird durch die Steuerzentrale festgestellt, daß eines zu einem Cluster  $i$  gehörendes Pixel am Videoeingang anliegt, so wird der entsprechende Schalter durch ein Signal  $I_i$  geschlossen und die zu diesem Pixel gehörende Helligkeitsinformation (Luminanzwert) wird dem zugeordneten Integrierer 16 zugeführt. Auf diese Weise liefert der Ausgang des Integrators nach Ablauf eines Vollbildes die Summenluminanz aller zum Cluster  $i$  gehörenden Pixelwerte. Die Ausgänge jeweils zweier nacheinander zu bewertender Cluster werden einem Komparator 17 zugeführt, der an seinem Ausgang eine logische 1 ausgibt, wenn der Luminanzwert des Clusters  $i$  größer ist als der des Clusters  $(i+1)$ . Sind alle für einen Vollbildzyklus definierten Cluster erfaßt und verglichen, speichert ein Taktimpuls  $Cl$  alle Ausgangswerte der Komparatoren 17 in Datenspeichern 18 und stellt deren Bits als Teil-Merkmalvektor (für ein Vollbild) einer Korrelationseinrichtung zur Verfügung. Jeder solche Teil-Merkmalvektor enthält die Daten eines Vollbildes.

Da jedem Vollbild  $n$  Cluster (z. B.  $n = 10$ ) entnommen werden, die  $n$ -te Differenz jedoch durch Vergleich mit dem ersten Cluster des nächsten Bildes zustandekommt, müssen  $n+2$  Halbleiterschalter und Integratoren vorhanden sein. Einer der beiden Integratoren  $16_{n+1}$  oder  $16_{n+2}$  speichert die Information des  $n$ -ten Clusters des vorhergehenden Bildes; der andere Integrator  $16_{n+2}$  oder  $16_{n+1}$  nimmt die Information des letzten Clusters des aktuellen Bildes auf und speichert sie für den Vergleich beim nächstfolgenden Bild. Für diese beiden Integratoren gilt eine besondere Regelung. Sie werden für jedes zweite Vollbild abwechselnd aktiviert. Nach Ablauf eines Bildes und vor Beginn des neuen Vollbildes werden alle Integratoren bis auf einen gelöscht, d. h. auf 0 zurückgesetzt. Im einfachsten Fall besteht ein Integrator aus einem RC-Glied, dessen Kondensator durch einen Reset-Impuls und einen elektronischen Schalter bzw. einen FET-Transistor entladen wird.

Zusätzlich stellt die Steuerzentrale 13 noch ein Signal zur Verfügung, das als "Vollbild-Flag" der Korrelationseinrichtung die Gültigkeit der in den Speichern 18 gespeicherten Werte anzeigt. Zur Speicherung dieses Flags ist ein weiterer Speicher 19 vorgesehen.

Alternativ könnte vorgesehen sein, nur  $n+1$  Halbleiterschalter und  $n+1$  Integratoren vorzusehen und die Information des  $n$ -ten Clusters abwechseln bei jedem Vollbild in den Integrator  $n$  oder den Integrator  $n+1$  einzuschreiben. Die Ansteuerung der entsprechenden Schalter  $15_n$  und  $15_{n+1}$  würde zusätzlich zu den Adressensignalen aus der Steuerzentrale 13 über einen Umschalter erfolgen, der durch einen bis zwei zählenden Binärzähler oder beispielsweise ein Flip-Flop angesteuert wird. Entsprechend müßte der letzte Komparator  $17_n$  der den Vergleich zwischen dem ersten Cluster des aktuellen Vollbildes mit dem  $n$ -ten Cluster des vorhergehenden Vollbildes durchführt, zwischen diesen beiden

Komparatoren umgeschaltet werden und auch der entsprechende Komparator  $17_n$  der den Vergleich zwischen dem  $n$ -ten Cluster und dem  $(n-1)$ -ten Cluster durchführt.

5 Noch eine weitere Variante der Erfindung besteht darin, daß der Vergleich nur innerhalb eines Vollbildes in der Weise zyklisch erfolgt, daß der  $n$ -te Cluster mit dem ersten Cluster desselben Vollbildes verglichen wird.

10 Fig. 6 zeigt eine zweite Variante für die Signalvorverarbeitung. Der Videoeingang 11, die Taktaufbereitung 12 und das Codierfeld 14 haben im Prinzip die gleiche Funktion wie im Ausführungsbeispiel der Fig. 5. Ähnliches gilt für die Steuerzentrale 13, die hier aber zusätzlich die Funktion als Adreßgenerator für einen RAM-Speicher (Random Access Memory) übernimmt. Die Taktaufbereitung 12 erzeugt dieselben Synchronsignale wie im Ausführungsbeispiel der Fig. 5 und führt diese der Steuerzentrale zu. Dort wird ein Bildpunkttakt (Pixeltakt ca. 12–14 Mhz) erzeugt, der einen Analog/Digitalwandler 20 veranlaßt, ca. alle 70 ns einen aktuellen Luminanzwert des Videosignales in ein 8-Bit-Wort umzuwandeln. Gleichzeitig erzeugt die Steuerzentrale 13 aus den Werten der Zeilen- und Spaltenzähler eine RAM-Adresse, unter der der aktuell gewandelte Luminanzwert kumulierend gespeichert wird. Hierzu wird einer arithmetisch-logischen Einheit (ALU) 21 zum einen der bisherige Speicherinhalt und zum anderen das Ergebnis des AD-Wandlers 20 zugeführt und der Ausgang, d. h. die Summe mit gleichem Takt unter der gleichen Adresse in einen der beiden RAM-Speicher 22 oder 23 abgespeichert. Jedem Cluster entspricht dabei eine Speicherzelle. Bei den Speichern handelt es sich um sogenannte Dual-Port-RAMs, die an ihrem anderen Port asynchron wieder ausgelesen werden können.

Ein Mikroprozessor 24 liest aus dem RAM 22 oder 23 die kumulierten Luminanzwerte aus, bildet die Differenzwerte zwischen "benachbarten" Clustern und gibt deren Vorzeichen an einen weiteren Speicher 25, der hier mit Ausgabelatch bezeichnet ist, weiter. An den Ausgängen dieses Ausgabelatch 25 stehen dann die gleichen Bits als Bild-Merkmalvektor an, wie an den Speicherplätzen  $18_1 - 18_n$  im Ausführungsbeispiel der Fig. 5.

Auch hier ist wie im Ausführungsbeispiel der Fig. 5 ein "Vollbildflag" vorgesehen, das aus einem Speicher 19 abzulesen ist.

Im Ausführungsbeispiel der Fig. 6 werden zwei identische Speicher 22 und 23 verwendet, wobei immer in einen Speicher das aktuelle Bild eingelesen wird, während das andere verarbeitet wird. Die Umschaltung erfolgt über die Steuerzentrale 13 durch entsprechende Ansteuerungen der Eingänge für Lesen (Rd), Schreiben (Wr) und Takt (Cl).

15 In einer praktischen Ausführungsform können die beiden Speicher 22 und 23 natürlich in einen integrierten Schaltkreis zusammengefaßt sein, ebenso wie auch die Funktion des Mikroprozessors 24 und der Steuerzentrale mit Adreßgenerator 13.

Das Codierfeld 14, das auch hier die Zuordnung der Pixel (Spalten- und Zeilenadresse) zu den einzelnen Clustern vornimmt, ist in einem Programmspeicher zu dem Mikroprozessor untergebracht. In Fig. 6 sind lediglich Funktionsblöcke zum besseren Verständnis getrennt gezeichnet.

20 Fig. 7 zeigt ein stark vereinfachtes Blockschaltbild einer Korrelationseinrichtung. Eine Signalvorverarbeitung 26 entsprechend den Ausführungsbeispielen der Fig. 5 oder 6 liefert jeweils bei Beendigung der Abta-



stung eines Vollbildes einen Teil-Merkmalsvektor, der beim hier beschriebenen Ausführungsbeispiel eine Länge von 10 Bit hat. Jeweils 51 aufeinanderfolgende Vollbilder ergeben einen Merkmalsvektor von 510 Bit. In einer "Lernphase" werden von Testmustern (ausgewählte Bildfolgen) entsprechende Merkmalsvektoren erzeugt und in einem Speicher 28 abgespeichert. Dies ist durch die gestrichelte Linie zwischen der Signalvorverarbeitung 26 und dem Speicher 28 angedeutet, sowie durch einen Schalter 70, der in der "Lernphase" geschlossen ist. Der Speicher enthält eine Anzahl  $m$  Merkmalsvektoren.

Während des normalen Betriebes erzeugt die Signalvorverarbeitung 26 ebenfalls alle 40 ms einen Teilmerkmalsvektor von 10 Bit Länge. Die Teilmerkmalsvektoren werden in ein Register 29 eingeschrieben, das insgesamt eine Anzahl  $n$  (beispielsweise 51) Teilbildvektoren enthält. Nach jedem Vollbild wird der letzte Teilbildvektor neu eingeschrieben und der erste Teilbildvektor gelöscht. Damit wird der Merkmalsvektor (mit einer Länge von 510 Bit) also laufend aktualisiert. Diese Aktualisierung erfolgt durch eine Steuerung 27, die von dem Speicher 19 der Fig. 5 oder 6 ein Signal für die Beendigung der Abtastung eines Vollbildes erhält. Sobald ein Teil-Merkmalvektor für ein neues Vollbild in das Register eingeschrieben ist, ruft die Steuerung 27 aus dem Speicher 28 nacheinander alle dort gespeicherten Merkmalsvektoren ab und schreibt diese zeitlich nacheinander, in ein Register 30. Die beiden Register 29 und 30 sind mit einem Vergleicher 31 verbunden, der die Registerinhalte vergleicht und bei einem vorgegebenen Grad der Übereinstimmung (z. B. 85 bis 95%) ein Signal "erkannt" an die Steuerung 27 abgibt. Innerhalb der 40 ms, die bis zum Eintreffen des nächsten Teil-Merkmalvektors des letzten Vollbildes verstreichen, vergleicht der Vergleicher alle vollständigen Merkmalsvektoren (Anzahl  $m$ ) die im Speicher 28 abgespeichert sind.

Werden die Registerinhalte der Register 29 und 30 als "übereinstimmend" erkannt, so kann die Steuerung 27 das weitere Auslesen von Merkmalsvektoren aus dem Speicher 28 abbrechen. Zusätzlich registriert die Steuerung 27 die Adresse desjenigen Merkmalsvektors im Speicher 28, bei dem Übereinstimmung auftrat und gibt ein entsprechendes Signal an eine nicht dargestellte Zentraleinheit ab.

Fig. 8 zeigt ein Prinzipschaltbild einer Korrelationsschaltung. Diese Korrelationsschaltung entspricht im wesentlichen der der EP 0 296 588, wo sie zur Wiedererkennung von Tonsignalen verwendet wird. Durch entsprechende Aufbereitung der Bildinformation mit den Ausführungsbeispielen der Fig. 1 bis 6 kann diese Schaltung auch für die Wiedererkennung von Bildsignalen verwendet werden. Die Ausgänge der Speicher 18<sub>1</sub>—18 Index  $n$  (Fig. 5) oder des Ausgabespeichers 25 (Fig. 6) werden über einen nicht dargestellten Parallel-Serien-Wandler als seriellles Signal  $D_1 N$  auf eine Leitung 37 gegeben, und zwar unter Steuerung eines Taktsignales  $\Phi$  das von einem nicht dargestellten Taktgeber auf einer Leitung 38 ankommt. Die seriell ankommenden Signale  $D_1 N$  werden zunächst in ein Schieberegister 39 eingeschrieben, wobei dieses Schieberegister vom Takt  $\Phi$  (Leitung 38) in Verbindung mit einem Gatter 40 gesteuert wird. Von diesem Schieberegister 39 gelangen die Signale  $D_1 N$  byteweise (8 Bit Länge) über einen Bustreiber 41 zu einem Speicher 42, der hier als RAM-Speicher mit zwei  $K \times 8$  Bit ausgelegt ist. Nach einer nicht dargestellten Variante ist es auch möglich, die an den Ausgängen der Speicher 18 bzw. 25 in paralleler Form vorlie-

genden Signale parallel in den Speicher 42 zu übernehmen, was durch das Vollbild-Flag des Speichers 19 gespeichert werden kann.

Eine Zählerkette mit hintereinander geschalteten Zählern 43, 44 und 45 erhält über ein UND-GATTER 46 das Taktsignal  $\Phi$  und bestimmt über seine Ausgänge die Adressen des Speichers 42, unter denen die vom Bustreiber 41 ankommenden Signale abgespeichert werden. Hierzu wird der Takt  $\Phi$  im Zähler 43 zunächst durch 8 geteilt und als Byteadresse an die Adreßleitungen A0 bis A10 des Speichers 42 gelegt.

Der 1/0-Übergang bestimmter Adreßleitungen steuert zwei mit negativen Flanken getriggerte JK-Flip-Flops 46 bzw. 47, welche wichtige Steuerfunktionen übernehmen. Bei einer Länge der Merkmalsvektoren von 512 Bit wird bei einem 1/0-Übergang der Adreßleitung A5 das erste Flip-Flop 46 getriggert, wenn ein Merkmalsvektor vollständig gespeichert ist. Dieses Flip-Flop wird zu Beginn einer "Lernphase" durch eine Taste 48 ("Lern-Startausführung") und zwei Gatter 49 und 50 gesetzt (Ausgang Q = 1; Q quer = 0) und kippt dann beim 1/0-Übergang der Adreßleitung A5 zurück und löscht die Zähler 43, 44 und 45. Es können dann weitere Merkmalsvektoren geladen werden, indem die beiden Zähler 44 und 45 über hier nicht gezeigte Leitungen auf eine Startadresse gesetzt werden. Nach jeweils weiteren 512 Takten wird wiederum über das Flip-Flop 46 der Vorgang gestoppt. Die mit "Maskenende" bezeichnete Leitung 51, die mit der Adreßleitung A5 verbunden ist, signalisiert einer hier nicht gezeigten Zentraleinheit das Ende des Einschreibevorganges einer Maske bzw. eines Merkmalsvektors, wodurch der Takt  $\Phi$  unterbrochen wird.

Wenn alle Merkmalsvektoren in den Speicher 42 eingelesen sind, wird ein Umschalter 70 in die dargestellte Stellung "K" für Korrelieren umgeschaltet. Gleichzeitig wird der Takt  $\Phi$  erhöht. Es werden dann die von den Signalvorverarbeitungen der Fig. 5 oder 6 ankommenden, auszuwertenden Signale des laufenden Programmes über eine Torschaltung 52 in nachfolgend näher beschriebener Weise in einen Speicher 53 eingelesen, und zwar wiederum mit derselben Taktrate, mit der die im Zwischenspeicher 42 abgespeicherten Test-Merkmalvektoren getaktet wurden. Während des Vergleichsvorganges werden nacheinander alle im Speicher 42 abgelegten Merkmalsvektoren mit dem 512 (bzw. 510) Bit des Speichers 53 verglichen, indem über zu diesen Speichern gehörende Zählerketten 43, 44, 45 zu Speicher 42; Zähler 54 und 55 zu Speicher 53, diese Zähler adreßrichtig ausgelesen werden. Der Zähler 43 ist nur "einkanalgig" belegt und enthält die letzten 512 Bit des an der Leitung 37 anliegenden Signales  $D_1 N$ . Der Speicher 42 gibt seine Daten über einen Multiplexer 56, der eine Parallel-Serienwandlung durchführt, an einen Eingang des Vergleichers ab, der hier als Exklusiv-Oder-Gatter 56 ausgeführt ist. Dem anderen Eingang dieses Exklusiv-Oder-Gatters 56 werden die Ausgangssignale des Speichers 53 zugeführt. Bei Übereinstimmung seiner beiden Eingangssignale (also 0 und 0 oder 1 und 1) erscheint am Ausgang des Exklusiv-Oder-Gatters 56 eine "1", welche über ein Gatter 57 einer Zählerkette aus den zwei 4-Bit-Zählern 58 und 59 zugeführt wird. Jede Übereinstimmung der Signale aus den beiden Speichern 42 und 53 erhöht also den Zählerstand der Zählerkette 58, 59 um eine 1.

Nach dem Vergleich jeweils eines Merkmalsvektors (512 Bit) wird durch ein Triggersignal (Adreßleitung A5) für das Flip-Flop 56 (Leitung "Maskenende") der Zähler-

stand der Zähler 58, 59 im D-Register 60, 61 übernommen und die Zähler 58, 59 anschließend gelöscht. Die D-Register 60, 61 geben ihre Daten an einen nicht-dargestellten Rechner weiter, was ebenfalls über das Signal auf der Leitung "Maskenende" (Leitung 62) gesteuert wird.

Bei vollständiger Übereinstimmung von Referenz-Merkmalvektor und Test-Merkmalvektor kann der Zählerstand der Zähler 58, 59 den Wert 512 enthalten. Wegen der fast immer vorhandenen Abtastfehler wird der Zählerstand in der Regler niedriger sein. Die Praxis zeigt, daß bei guter Signalaufbereitung etwa 85–95% des maximalen Zählerstandes bei Koinzidenz erreicht wird (ein Zählerstand von 256 ist der Erwartungswert bei nicht-korrelierten Signalen). Wenn alle Merkmalsvektoren des Speichers 42 mit den aktuell im Speicher 53 gespeicherten 512 Bit verglichen sind, wird dies — je nach Anzahl der Merkmalsvektoren — auf einem der Ausgänge der Zähler 44 bzw. 45 erkannt und über eine Leitung 63 dem Flip-Flop 47 gemeldet, welches umschaltet und dafür sorgt, daß ein Taktimpuls für die Zählerkette 43, 44, 45 unterdrückt wird, während die Zählerkette 54, 55 des Speichers 53 weiterschaltet und über einen Impuls auf der Leitung 64 (Ausgang des Flip-Flop 47) das Tor 52 für einen Taktimpuls öffnet und damit einen Abtastwert des Signales  $D_{IN}$  auf der Leitung 37 in den Speicher 53 einschreibt. Dabei schaltet die Zählerkette 54, 55 auch um einen Takt weiter, so daß zwischen den Adressen der beiden Speicher sich eine bei Vergleich jedes vollständigen Merkmalsvektor eine um eine 1 vergrößernde Differenz ergibt. Damit ist der Speicher 53 relativ zum Speicher 42 wie ein Schieberegister organisiert, welches nach einem Durchlaufen bzw. Vergleichen aller Merkmalsvektoren des Speichers 42 um einen Schritt weiter schaltet.

Sobald das ein Durchlaufen aller Merkmalsvektoren anzeigende Signal auf der Leitung 63 verschwindet, wird das Flip-Flop 47 über das nächste Taktsignal  $\Phi$ , welches über Gatter 64, 65 und 66 zu dessen Rücksetzeingang R gelangt, wieder rückgesetzt, wodurch das Gatter 46, welches die Taktimpulse zu der Zählerkette 43, 44, 45 läßt, wieder freigegeben wird. Kurz zusammengefaßt werden also innerhalb von 40 ms, d. h. der Zeitdauer zum Abtasten eines Vollbildes des Videosignales alle gespeicherten Merkmalsvektoren zu je 512 Bit verglichen. Das Flip-Flop 46 überwacht dabei den einzelnen Merkmalsvektor (512 Bit), während das Flip-Flop 47 die Gesamtzahl aller gespeicherten Merkmalsvektoren überwacht. Je nach Anzahl der im Speicher 42 gespeicherten Merkmalsvektoren ist der Takteingang (Leitung 63) des Flip-Flop 47 mit einem anderen Ausgang der Zählerkette 44, 45 verbunden.

Im beschriebenen Ausführungsbeispiel wird der Vergleich seriell mit einem Exklusiv-Oder-Gatter 56 durchgeführt. Da die Taktfrequenz für den Vergleich mit der Anzahl der gespeicherten Merkmalsvektoren steigt, ist die Anzahl der Merkmalsvektoren durch die maximale Arbeitsfrequenz der digitalen Bausteine begrenzt. Es ist daher prinzipiell auch möglich, mehrere parallel arbeitende Vergleiche einzusetzen, sei es, daß der Vergleich byteweise (8 Bit) durchgeführt wird, sei es auch, daß im Extremfall alle 512 Bits parallel in 512 parallelen EX-OR-Gattern verglichen würden. Letzteres ist allerdings relativ aufwendig, da Speicher mit 512 Ausgängen benötigt würden.

Weiterhin ist es prinzipiell auch möglich, mehrkanalig zu arbeiten, dergestalt, daß zwei oder mehrere Sender überwacht werden. Hierzu könnte man entweder den

Speicher 53 vergrößern und derart organisieren, daß zwei verschiedene Eingangssignale  $D_{IN}$  parallel bzw. zeitlich verschachtelt eingelesen und verglichen werden oder daß statt des einen Speichers 53 pro Kanal ein eigener entsprechender Speicher vorgesehen ist.

Schließlich ist es auch möglich, den Speicher 42 als Datenbank zu organisieren und über einen Suchalgorithmus zu untersuchen, ob der im Speicher 53 vorhandene "Datensatz" in der "Datenbank" des Speichers 42 vorhanden ist. Der Suchalgorithmus kann dabei auch so aufgebaut sein, daß nicht nur nach "identischen" Datensätzen, sondern nach "ähnlichen" Datensätzen (Datensatz = ein Merkmalsvektor) gesucht wird.

Da es bei der Wiedererkennung von Bildern bzw. Bildfolgen nicht nur darauf ankommt, ob eine der zu überwachenden Bildfolgen gespielt wurde, sondern auch darauf, welche von diesen, kann gleichzeitig mit dem Korrelationsergebnis der Zählerkette 58, 59 der Zählerstand der Zählerkette 43, 44, 45 abgefragt und in die nicht dargestellte Zentraleinheit eingegeben werden. Damit ist es möglich, den aktuell erkannten Merkmalsvektor (Bildfolge) zu identifizieren und dieses Identifikationsergebnis zusammen mit Datum und Uhrzeit und ggf. auch den überwachten Sendekanal zu speichern oder auszugeben.

Zu den Speichern 42 und 53 ist noch anzumerken, daß die mit WE/quer bzw. OE/quer bezeichneten Eingänge Steuereingänge sind, wobei der Eingang WE/quer ein negierter "write enable"-Eingang ist, der das Einschreiben in den Speicher steuert, während der Eingang OE/quer ein negierter "output-enable"-Eingang ist, der das Auslesen von Daten aus dem Speicher steuert. Durch die diesen Eingängen vorgeschalteten Gatter wird die Verbindung mit dem Taktsignal  $\Phi$  und dem Signal auf der Leitung 64 bezüglich des Speichers 53 bzw. durch die Stellung des Schalters 70 und das die Übertragung eines Bytes (8 Bit) aus dem Speicher 42 identifizierende Signal des Zählers 43 für den Speicher 42 überwacht, daß die beiden Speicher 42 und 53 nur zu den jeweils richtigen Zeitpunkten ausgelesen bzw. eingeschrieben werden können. Die übrigen in der Zeichnung dargestellten, jedoch in der Beschreibung nicht ausdrücklich erwähnten Gatter bzw. deren externe Beschaltungen dienen ebenfalls der richtigen Funktion der Ablaufsteuerung, wobei deren Funktion dem Fachmann ohne weiteres ersichtlich ist, bzw. sich beim Aufbau der dargestellten Schaltung von selbst einstellt.

#### Patentansprüche

1. Verfahren zum Wiedererkennen von unikaten Bildsignalen, gekennzeichnet durch folgende Schritte:

- a) Aus Helligkeitswerten einer Anzahl  $n$  von Pixeln an mehreren vorbestimmten Stellen jedes Vollbildes wird je ein erstes Helligkeitssignal ermittelt,
- b) diese ersten Helligkeitssignale von in einer vorgegebenen Reihenfolge aufeinanderfolgenden Stellen werden voneinander subtrahiert,
- c) die Vorzeichen dieser Differenzen werden je als 1 Bit digitalisiert (Signumfunktion),
- d) die digitalisierten Signale werden als Test-Bitfolge (Merkmalsvektor) abgespeichert,
- e) die Schritte a) bis d) werden für eine Anzahl  $m$  (z. B. 51) aufeinanderfolgende Vollbilder wiederholt,



f) zu untersuchende Bildsignale werden in gleicher Weise (Schritte a) bis e)) zur Erzeugung einer Referenz-Bitfolge (Referenz-Merkmalvektor) aufbereitet und

g) jede Test-Bitfolge wird mit allen Referenz-Bitfolgen verglichen und bei Überschreiten eines vorgegebenen Maßes der Übereinstimmung wird ein Signal für "Bildfolge" wiedererkannt erzeugt.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß in Schritt a) räumlich zusammenhängende Pixel eines Bildes einen Cluster bilden, dessen Helligkeitssignal ermittelt wird.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die Pixel jedes Clusters in Form eines Quadrates unmittelbar benachbarter Pixel angeordnet sind.

4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, daß jeder Cluster  $5 \times 5$  Pixel enthält.

5. Verfahren nach einem der Ansprüche 2 bis 4, dadurch gekennzeichnet, daß die Helligkeitsswerte (Luminanzwerte) aller Pixel eines Clusters zu dem Helligkeitssignal aufsummiert werden.

6. Verfahren nach einem der Ansprüche 2 bis 4, dadurch gekennzeichnet, daß der arithmetische Mittelwert der Helligkeitsswerte aller Pixel eines Clusters das Helligkeitssignal bildet.

7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß das Subtrahieren der Helligkeitssignale von in einer vorgegebenen Reihenfolge benachbarten Stellen und das Digitalisieren der Vorzeichen dieser Differenzen in einem Arbeitsschritt mittels eines Komparators durchgeführt wird.

8. Verfahren nach einem der Ansprüche 2 bis 7, dadurch gekennzeichnet, daß die Cluster statistisch über die Fläche des Vollbildes verteilt sind.

9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß die Reihenfolge, in der die Helligkeitssignale benachbarter Cluster voneinander subtrahiert bzw. miteinander verglichen werden, so festgelegt ist, daß der Mittelwert der räumlichen Abstände der in dieser Reihenfolge jeweils benachbarten Cluster einen Maximalwert einnimmt.

10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß der in der Reihenfolge letzte Cluster mit dem ersten Cluster des nächsten Vollbildes verglichen wird.

11. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß der in der Reihenfolge letzte Cluster mit dem ersten Cluster des nächstfolgenden Vollbildes verglichen wird.

12. Verfahren nach einem der Ansprüche 8 bis 11, dadurch gekennzeichnet, daß mindestens zwei voneinander verschiedene Clusteranordnungen (Masken) vorgesehen sind und daß aufeinanderfolgende Vollbilder mit zyklisch abwechselnder Clusteranordnung (Masken) abgetastet werden.

13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß bei Verwendung von zwei Clusteranordnungen (Masken) diese einander gleich jedoch gegenüber einem abzutastenden Bild um 90 Grad gegeneinander gedreht sind.

14. Verfahren nach Anspruch 12 oder 13, dadurch gekennzeichnet, daß jede Testbildfolge mit allen zyklischen Clusterfolgen (Masken) abgetastet wird, daß jede Referenzbildfolge nur mit einer zyklischen Clusterfolge (Maske) abgetastet wird und daß jede

so ermittelte Referenzbitfolge nur mit einer ausgewählten der verschiedenen Testbildfolgen korreliert wird.

15. Verfahren nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß der Vergleich jeder Test-Bitfolge mit allen Referenzbitfolgen jedesmal dann stattfindet, wenn ein Vollbild abgetastet ist.

16. Vorrichtung zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 15, gekennzeichnet durch:

- einen Taktaufbereitungsschaltkreis (12), der aus einem Videosignal mindestens ein Zeilentaktsignal und ein Spaltentaktsignal erzeugt,
- ein Codierfeld (14), das vorbestimmte Pixel des im Videosignal enthaltenen Bildes bezeichnet,

- eine Steuereinrichtung (13), die mit dem Taktaufbereitungsschaltkreis (12) und dem Codierfeld (14) verbunden ist und mindestens einen Summierer ( $21; 16_1 - 16_{n+2}$ ), so ansteuert, daß Luminanzwerte des Videosignales, der von dem Codierfeld (14) vorbestimmten und durch das Zeilentaktsignal und das Spaltentaktsignal identifizierten Pixel in zu Clustern zusammengefaßten Gruppen aufsummiert werden,

- eine Vergleichseinrichtung ( $24; 17_1 - 17_n$ ) in der die Summensignale der einzelnen Cluster mit Summensignalen anderer vorbestimmter Cluster verglichen werden und pro Vergleich ein digitales Signal in Form eines Bits erzeugt wird, das anzeigt, welcher der beiden verglichenen Cluster einen größeren Luminanzwert hat,

- einen Speicher ( $25; 18_1, 18_n$ ), in dem die Vergleichsergebnisse gespeichert werden und

- eine Korreliereinrichtung (Fig. 7), in welcher der Inhalt des Speichers ( $25; 18_1 \dots 18_n$ ) mit in einem weiteren Speicher ( $30; 42$ ) gespeicherten Testbildmustern in einer vorbestimmten Taktfolge laufend korreliert werden.

17. Vorrichtung nach Anspruch 16, dadurch gekennzeichnet, daß die Summierer Integratoren ( $16_1 - 16_{n+2}$ ) sind, die über von der Steuereinrichtung (13) angesteuerte Schalter ( $15_1 - 15_{n+2}$ ) mit dem Videosignal beaufschlagt werden.

18. Vorrichtung nach Anspruch 16, dadurch gekennzeichnet, daß der Summierer eine arithmetisch/logische Recheneinheit (21) ist, deren einem Eingang über einen Analog/Digital-Wandler (20) ein digitalisiertes Videosignal und deren anderem Eingang ein in einem Zwischenspeicher (22, 23) zwischengespeicherter Wert zugeführt wird und deren Ausgang mit dem Zwischenspeicher (22, 23) verbunden ist.

19. Vorrichtung nach Anspruch 17, dadurch gekennzeichnet, daß die Ausgänge der Integratoren ( $16_1 - 16_{n+2}$ ) jeweils paarweise mit einem Komparator ( $17_1 - 17_n$ ) verbunden sind, der die in den zugeordneten Integratoren aufsummierten Luminanzsignale vergleicht und in Abhängigkeit von diesem Vergleich ein Signal an den Speicher ( $18_1 - 18_n$ ) ausgibt.

20. Vorrichtung nach Anspruch 19, dadurch gekennzeichnet, daß einem Eingang der Komparatoren ( $17_1 - 17_n$ ) zusätzlich ein vorbestimmter Schwellwert hinzuaddiert wird.

21. Vorrichtung nach Anspruch 18, dadurch ge-

kennzeichnet, daß die in den Zwischenspeichern (22, 23) für die einzelnen Cluster abgespeicherten Summensignale unter Steuerung eines Mikroprozessors (24) in vorbestimmter Reihenfolge miteinander verglichen werden und das Vergleichsergebnis in den Speicher (25) abgelegt wird. 5

Hierzu 8 Seite(n) Zeichnungen

10

15

20

25

30

35

40

45

50

55

60

65

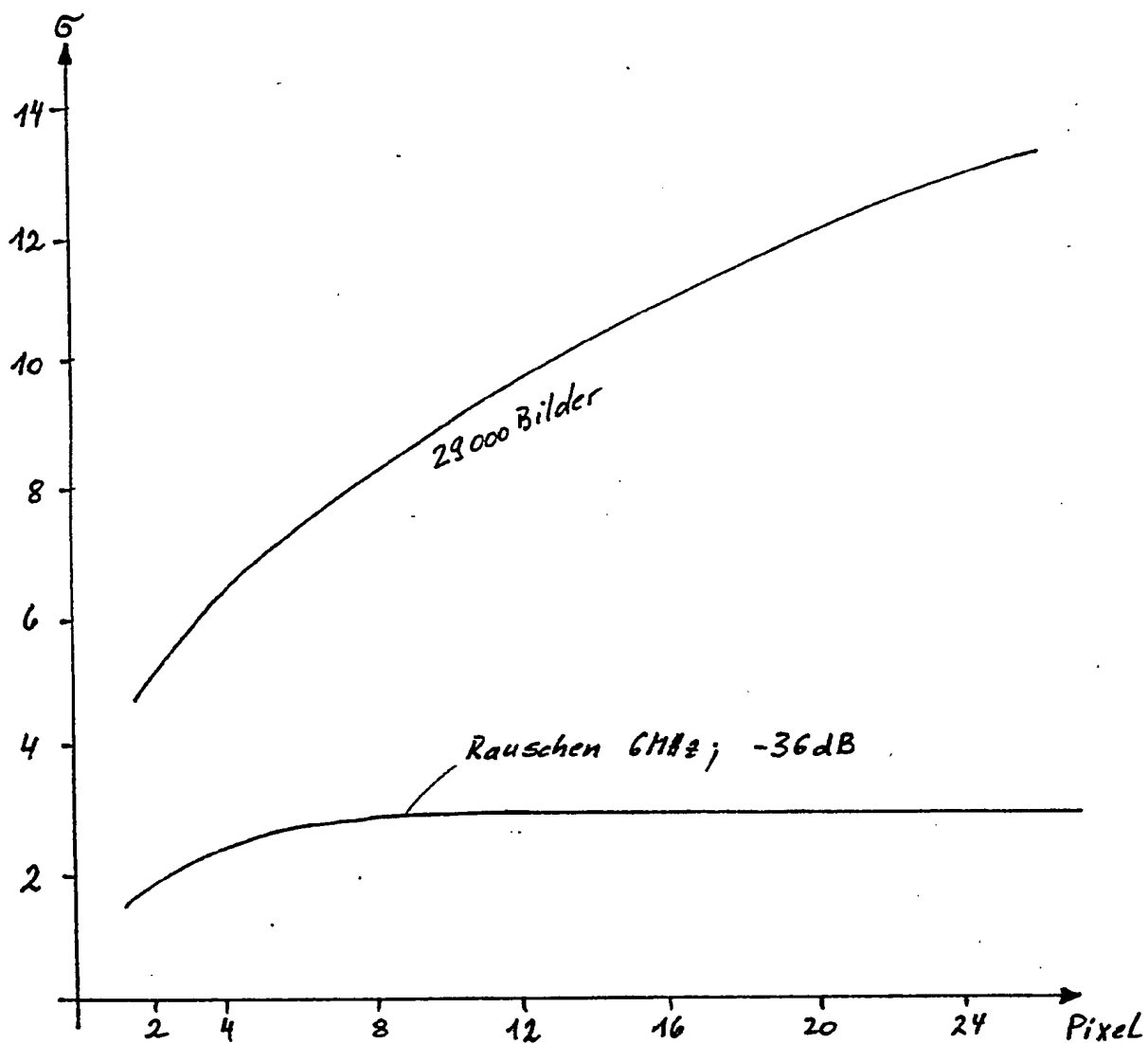


Fig. 1

## Referenzmuster

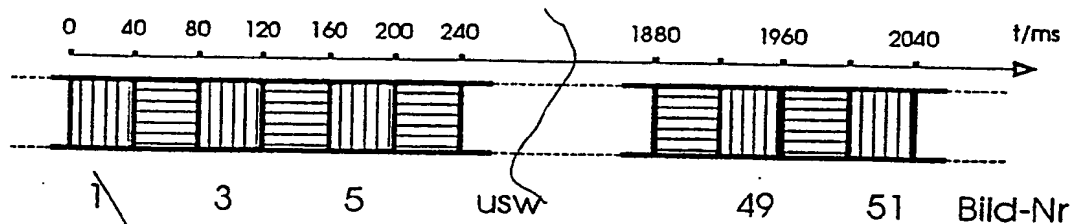


Fig. 2c

10 Punkte je Bild mit 10 Bit codiert; 51 Bilder ergeben 510 Bit

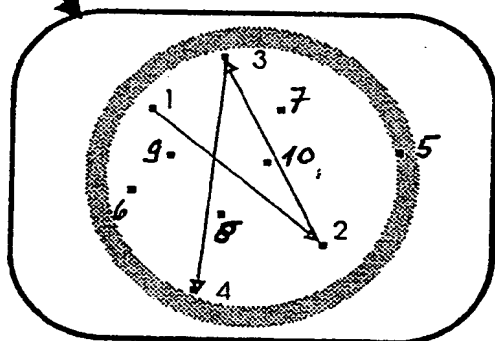


Fig. 2a

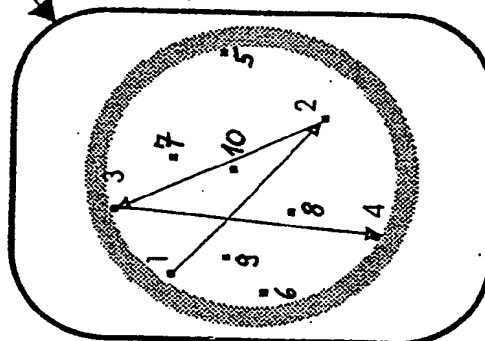
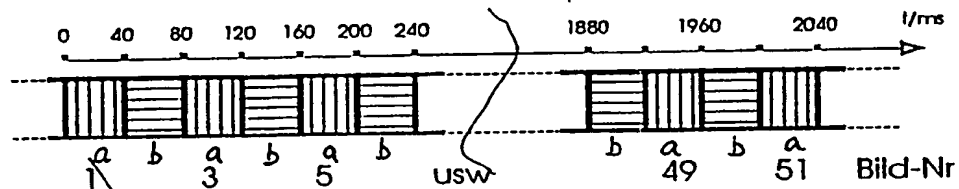


Fig. 2b

Fig. 2



10 Punkte je Bild mit 10 Bit codiert; 51 Bilder ergeben 510 Bit

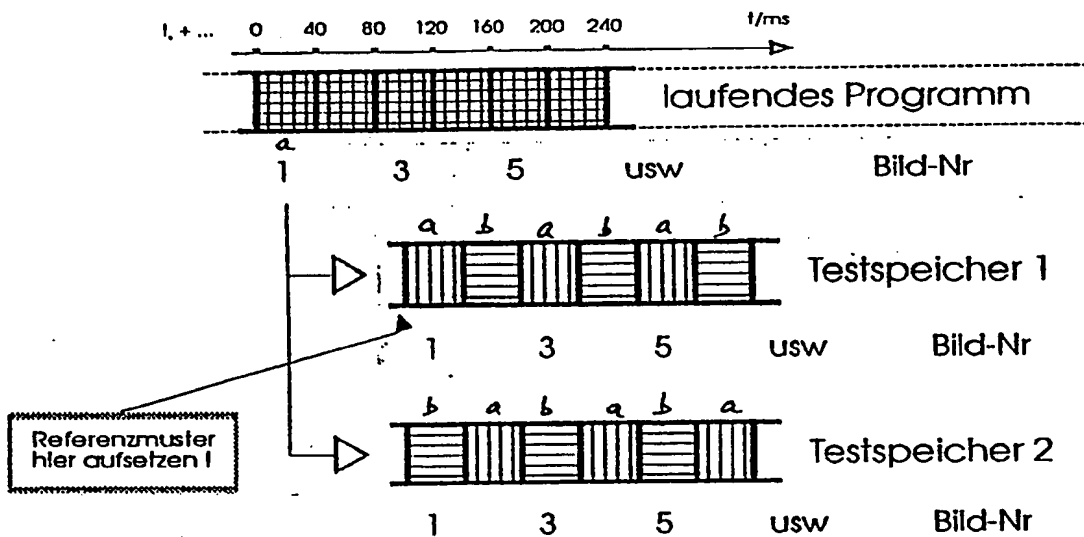


Fig. 3

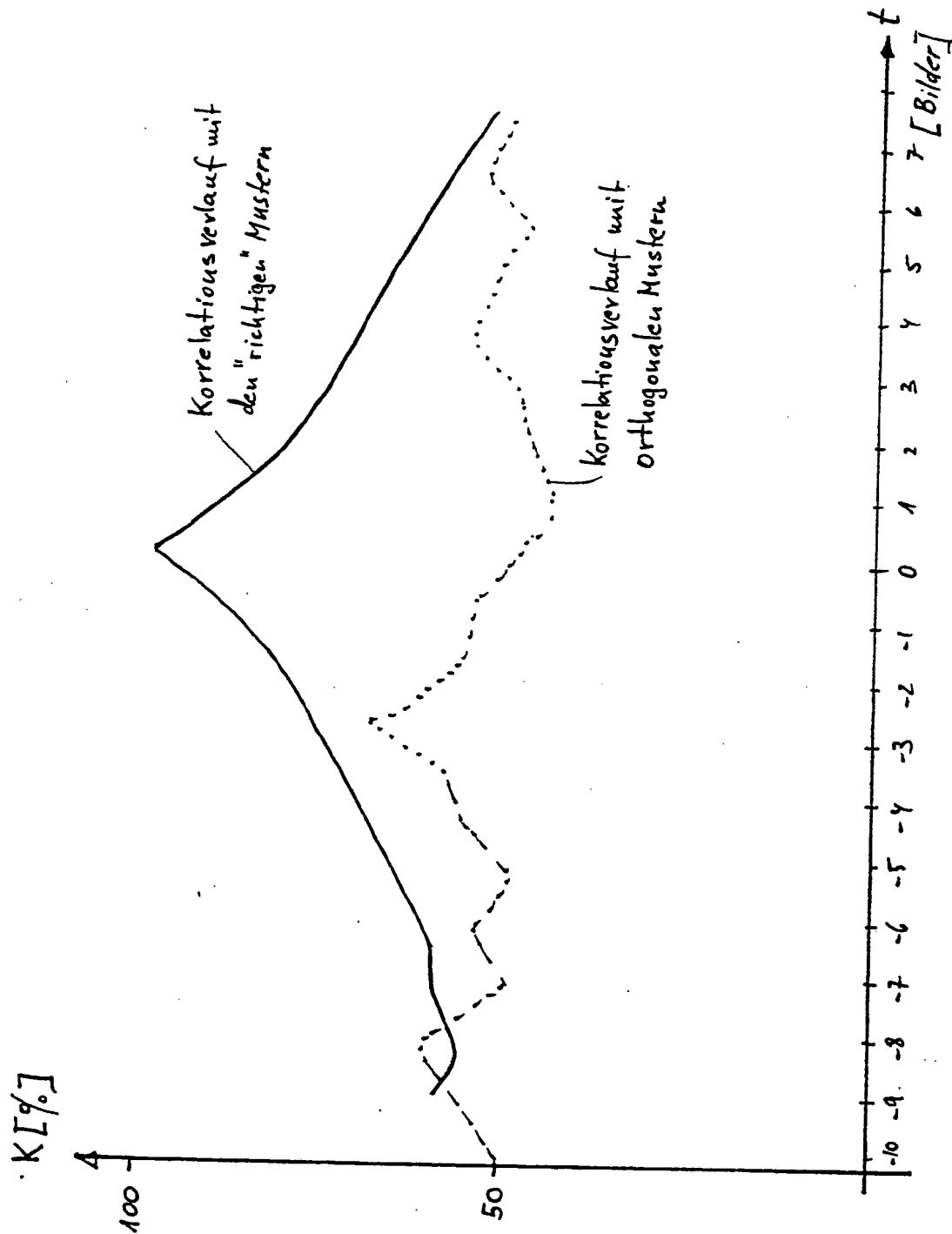


Fig. 4



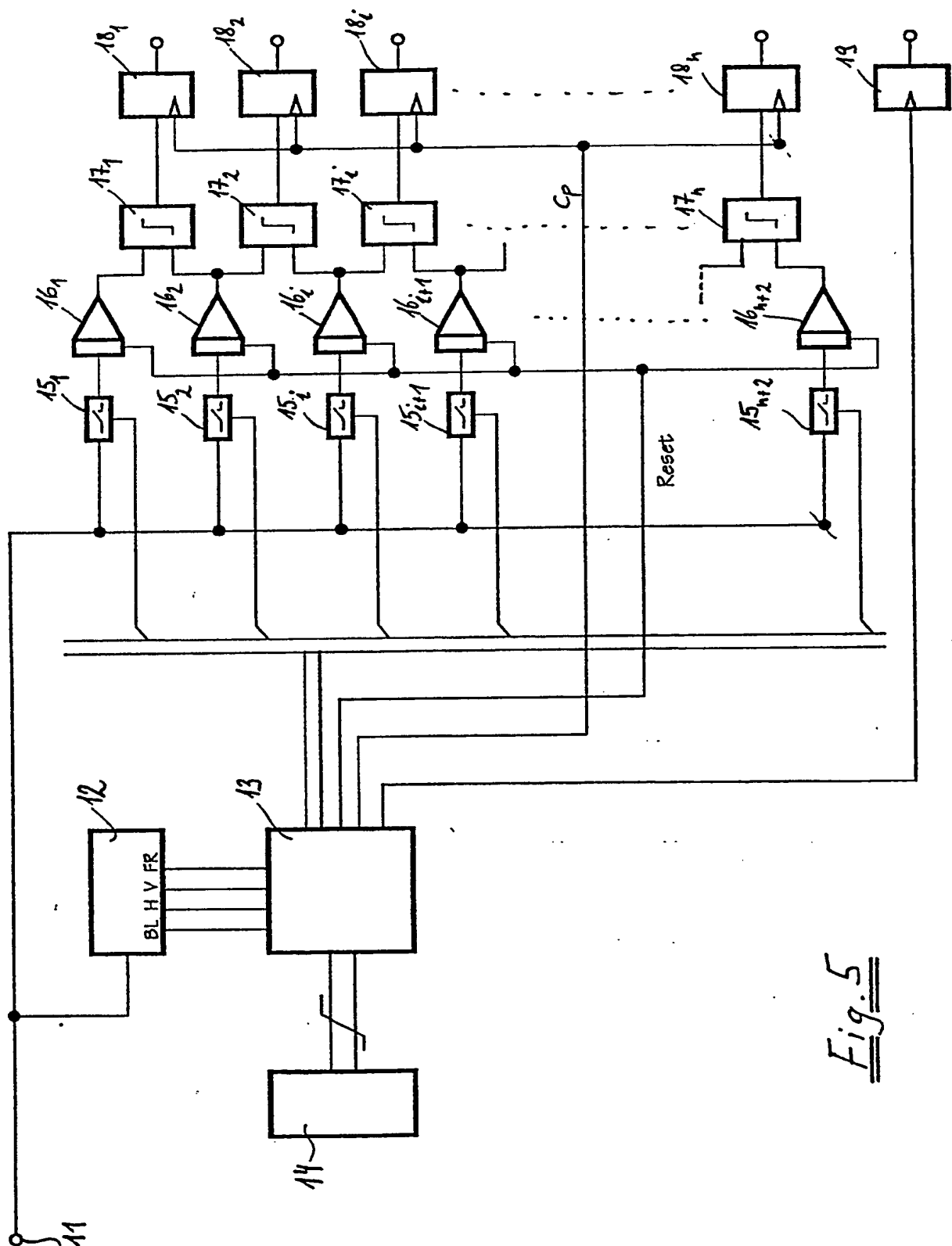


Fig. 5

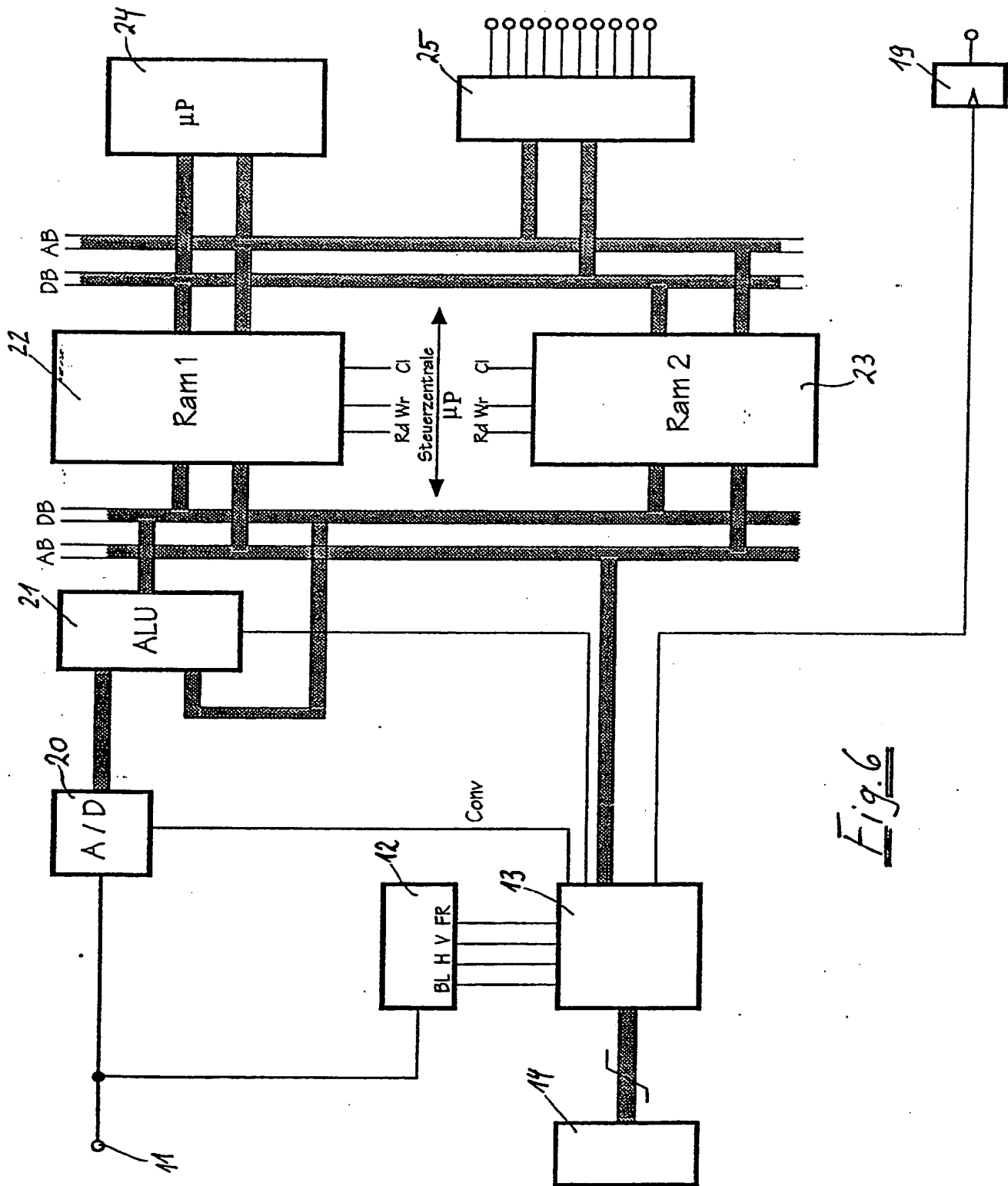


Fig. 6

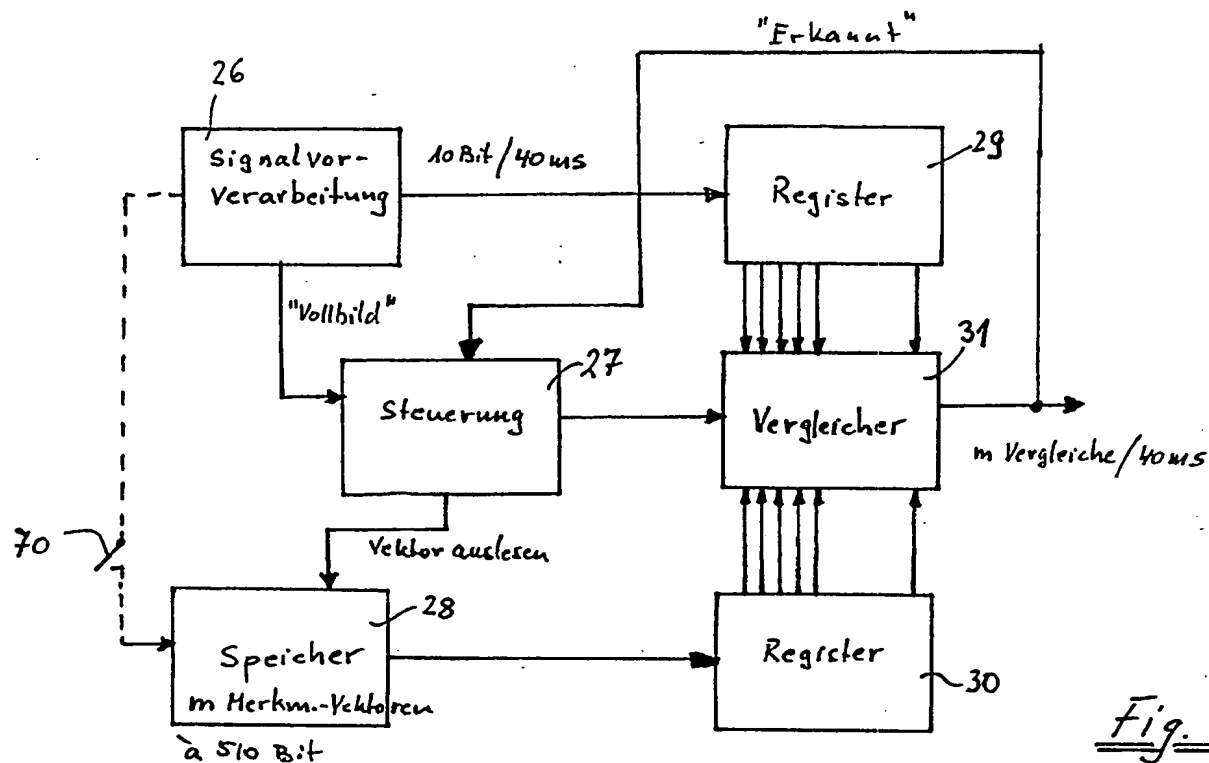


Fig. 8

